12-2017

Exploring spatio-temporal effects in traffic crash trend analysis

Chenhui Liu
Iowa State University, cliu9@iastate.edu

Anuj Sharma
Iowa State University, anujs@iastate.edu

Exploring spatio-temporal effects in traffic crash trend analysis

Abstract

Unobserved heterogeneity produced by spatial and temporal correlations of crashes often needs to be captured in crash frequency modeling. Although many studies have included either spatial or temporal effects in crash frequency modeling, only a limited number of studies have considered both. This study addresses the limitations of existing studies by exploring multiple models that best fit the spatial and temporal correlations. In this study, we used Bayesian spatio-temporal models to investigate regional crash frequency trends, and explored the effects of omitting spatial or temporal trends in spatio-temporal correlated data. The fast Bayesian inference approach, integrated nested Laplace approximation, was used to estimate parameters. It was found that fatal crashes showed decreasing trends in all Iowa counties from 2006 to 2015, but the decreasing rates varied by counties. Among all the covariates investigated, only vehicle miles traveled (VMT) was significant. None of the socio-economic or weather indicators were found to be significant in the presence of VMT. Both spatial and temporal effects were found to be important, and they were responsible for both over dispersion and zero inflation in the crash data. In addition, spatial effects played a more important role than did temporal effects in the studied dataset, but temporal component selection was still important in spatio-temporal modeling.

Keywords

Spatio-temporal modeling, Bayesian, Integrated nested Laplace approximation, Conditional autoregressive, Unobserved heterogeneity

Disciplines

Civil Engineering | Multivariate Analysis | Transportation Engineering

Comments

This is a manuscript of an article published as Liu, Chenhui, and Anuj Sharma. "Exploring spatio-temporal effects in traffic crash trend analysis." *Analytic Methods in Accident Research* 16 (2017): 104-116. DOI: 10.1016/j.amar.2017.09.002. Posted with permission.

Exploring Spatio-Temporal Effects in Traffic Crash Trend Analysis

Chenhui Liu[ab*], Anuj Sharma[a]

[a]Department of Civil, Construction, and Environmental Engineering, Iowa State University.

InTrans, 2711 South Loop Drive, Suite 4700, Ames, IA 50010-8664

[b]Department of Statistics, Iowa State University

*Corresponding author

Email addresses: cliu9@iastate.edu (C. Liu), anujs@iastate.edu (A. Sharma)

Abstract

Unobserved heterogeneity produced by spatial and temporal correlations of crashes often needs to be captured in crash frequency modeling. Although many studies have included either spatial or temporal effects in crash frequency modeling, only a limited number of studies have considered both. This study addresses the limitations of existing studies by exploring multiple models that best fit the spatial and temporal correlations. In this study, we used Bayesian spatio-temporal models to investigate regional crash frequency trends, and explored the effects of omitting spatial or temporal trends in spatio-temporal correlated data. The fast Bayesian inference approach, integrated nested Laplace approximation, was used to estimate parameters. It was found that fatal crashes showed decreasing trends in all Iowa counties from 2006 to 2015, but the decreasing rates varied by counties. Among all the covariates investigated, only vehicle miles traveled (VMT) was significant. None of the socio-economic or weather indicators were found to be significant in the presence of VMT. Both spatial and temporal effects were found to be important, and they were responsible for both over dispersion and zero inflation in the crash data. In addition, spatial effects played a more important role than did temporal effects in the studied dataset, but temporal component selection was still important in spatio-temporal modeling.

Keywords: spatio-temporal modeling, Bayesian, Integrated Nested Laplace Approximation, conditional autoregressive, unobserved heterogeneity

1 Introduction

Traffic crashes have been one of the major sources of fatalities and injuries in the United States. Crash frequency models often are used to identify the factors influencing the propensity of traffic crashes. The most common crash frequency model is the Poisson model. When crashes show over dispersion, quasi-Poisson, Poisson log-normal model (PLN), and negative binomial (NB) models are often adopted. Unobserved heterogeneity is often an issue in crash frequency analysis, because many crash-related factors are often not observed by the analyst (Mannering et al., 2016). The excess zeros in crash data can be a result of unobserved heterogeneity (Mullahy, 1997), often causing zero-inflated and hurdle models to be adopted (Lord et al., 2005; Lord and Mannering, 2010; Malyshkina and Mannering, 2010; Mannering et al., 2016; Mannering and Bhat, 2014). In addition, the zero-state Markov switching model, which allows observations to switch between zero and normal-count states over time, has been proven to be a viable alternative to zero-inflated models (Malyshkina and Mannering, 2010). Because crash data are often aggregated over time and space, spatial and temporal correlations are often also responsible for a portion of unobserved heterogeneity, as crashes that occur close in space or time are very likely to share some unobserved characteristics (Lord et al., 2005; Lord and Mannering, 2010; Mannering et al., 2016; Mannering and Bhat, 2014; Savolainen et al., 2011). However, these spatial and temporal correlations are often overlooked in existing studies, and neglecting them may produce inefficient or biased estimated results (Mannering et al., 2016; Mannering and Bhat, 2014; Savolainen et al., 2011).

The spatial correlation of traffic crashes may exist on a macro- or microscopic spatial scale. At a macroscopic level, factors such as census tract (Wang and Kockelman, 2013), traffic analysis zone (Matkan and Mohaymany, 2013), ZIP code level (Ponicki et al., 2013), census block group (Noland et al., 2013), census ward (Boulieri et al., 2016; Quddus, 2008), county (Aguero-Valverde and Jovanis, 2006; Eckley and Curtin, 2013; Song et al., 2006), and state/province (Erdogan, 2009; Truong et al., 2016), as well as similarity of economic and social activities, culture, land use, and enforcements within a given region, may explain the spatial correlation in traffic crashes. At a microscopic level, crashes occurring at nearby intersections (Abdel-Aty and Wang, 2006; Ahmed and Abdel-Aty, 2015; Guo et al., 2010; Liu et al., 2015; Mitra et al., 2007; Pulugurtha and Sambhara, 2011; Wang and Abdel-Aty, 2006; Xie et al., 2014) or adjacent road segments (Aguero-Valverde, 2011; Aguero-Valverde and Jovanis, 2008; Jiang et al., 2014; Wang et al., 2011, 2009; Zeng and Huang, 2014) may be correlated as a result of geometric or traffic flow similarities (Levine et al., 1995).

Temporal correlation captures the variability of traffic crashes with temporal scales such as year (Andrey, 2010; Boulieri et al., 2017; Brijs et al., 2008; El-Basyouny et al., 2014; Malyshkina and Mannering, 2010; Matkan and Mohaymany, 2013; Wang et al., 2011; Wang and Abdel-Aty, 2006; Yannis et al., 2011), month (Hu et al., 2013; Quddus, 2008), week (Kilamanua et al., 2011; Liu et al., 2015; Malyshkina et al., 2009; Sukhai et al., 2011), day (Brijs et al., 2008), and hour (Kilamanua et al., 2011; Liu et al., 2015). Temporal correlation reflects the influence of different traffic-related factors, such as economy, weather, environment, law, and travel demand, which often exhibit some temporal trends or periodicities.

Depending on the study site, one of three scenarios is feasible: (a) the crash data may show both spatial and temporal effects, (b) these effects may exist individually, or (c) neither of them may exist. When spatial and temporal effects co-exist, their interaction (i.e. spatio-temporal effects) also needs to be considered. Although many studies have included either spatial effects or temporal effects in crash frequency modeling, only a limited number of studies have considered both of them. Miaou et al. (2003) first introduced the spatio-temporal modeling approach to traffic crash modeling in analyzing yearly county-level crash rates in Texas from 1992 to 1999 using multiple spatio-temporal models. Wang and Abdel-Aty (2006) analyzed spatial and temporal correlations for rear-end crashes at signalized intersections in Florida. However, they built separate models for spatial effects and temporal effects. Jiang et al. (2014) considered both spatial and temporal correlations in analyzing the crashes on urban four-lane divided arterial segments in the central Florida area. However, they assumed that the spatial and temporal effects followed normal distributions without presenting any data-driven evidence to support their assumption. Truong et al. (2016) analyzed yearly crash fatalities of 63 provinces in Vietnam from 2012 to 2014 using the conditional autoregressive (CAR) spatio-temporal autocorrelation technique. The CAR spatio-temporal model performed better than the random effects NB model and random parameters NB model did in terms of both goodness of fit and crash prediction. Aguero-Valverde and Jovanis (2006) had similar findings.

The CAR model (Besag, 1974; Besag et al., 1991) often is used for modeling areal data in spatial statistics. Several researchers (Aguero-Valverde and Jovanis, 2006; Boulieri et al., 2017; Truong et al., 2016; Wang et al., 2011) have used the CAR model to illustrate spatial correlations paired with different temporal models. However, they all showed only one temporal model, despite the fact that the choice of a particular temporal model was also very important (Miaou et al., 2003). In this study, we used the spatio-temporal crash frequency model to identify the long-term county-level fatal crash frequency trends in Iowa. Multiple temporal components were built and contrasted to choose the most appropriate model. A fast Bayesian estimation tool, integrated nested Laplace approximation (INLA), was used to estimate these spatio-temporal models.

The workflow of the data analysis is as follows:

- First, we discuss whether crashes have over dispersion and zero inflation.
- Second, we examine spatial correlations and temporal correlations of crashes.
- Third, we evaluate the necessity of including the spatial component, temporal component, and spatio-temporal component in modeling, and we also discuss the temporal component selection.
- Finally, after determining the final model, the estimation results are discussed.

The rest of paper is organized as follows. Section 2 comprises a discussion of the traffic crash data used for this study. Section 3 presents the statistical models and estimation methods used in this study. Section 4 includes the analyses and discussions of the observed results. A conclusion and future recommendations are provided in section 5.

2 Data Description

Traffic crash data for Iowa's 99 counties from 2006 to 2015 were obtained from the Iowa Department of Transportation. Based on their severity, the crashes were divided into five categories: fatal, major injury, minor injury, possible injury/unknown, and property damage only. Fatal crashes were analyzed for this study, as they usually cause much more severe outcomes than do other types of crashes. The vehicle miles traveled (VMT) data for each county in each year from 2006 to 2015 were downloaded from the website of the Iowa Department of Transportation (2016). In addition, population and unemployment rate data were downloaded from the website of Iowa Community Indicators Program (2016), and per capita personal income data were downloaded from the website of the U.S. Bureau of Economic Analysis (2016). Because weather has been shown to significantly influence crash frequencies in many studies (Brijs et al., 2008; Golob and Recker, 2003; Knapp et al., 2000; Maze et al., 2005), rainfall amounts, snowfall amounts, and the number of days with a minimum temperature higher than 32°F (TH32) were downloaded from the website of the Iowa Environmental Mesonet (2017). These weather data were collected based on the daily climate observations from the National Weather Service's Cooperative Observer Program. A summary of the variables is given in Table 1.

The variance of fatal crashes was larger than the mean, which implied that over-dispersion was occurring. The proportion of zero crashes, used to preliminarily check whether or not zero-inflated models are needed, is shown in the last column of Table 1. The zero proportion of fatal crashes was 0.113, much larger than 0.034, which was the supposed probability value of zero under a Poisson distribution with the mean being 3.383. This implies that the zero-inflated model may be considered. Additionally, the highest correlation among the covariates was –0.338 (between TH32 and snowfall). Thus, no explanatory variables showed strong correlations in this study.

Table 1 Descriptive statistics of collected variables

Variables	Mean	Std. Error	Median	Min.	Max.	Zero-proportion
Fatal crash frequency	3.383	3.818	2.000	0.000	35.000	0.113
VMT (1,000,000 miles)	0.320	0.487	0.186	0.047	4.215	—
Population (10,000)	3.076	5.273	1.571	0.380	46.771	—
Unemployment rate (%)	4.846	1.347	4.600	2.000	10.200	—
Income ($10,000)	3.877	0.666	3.877	2.247	6.464	—
Rainfall (inch)	38.390	8.570	38.610	17.850	64.990	—
Snowfall (inch)	34.560	14.377	35.000	0.000	85.100	—
TH32 (days)	222.600	15.733	221.000	174.000	272.000	—

Note: VMT, vehicle miles traveled; TH32, number of days with minimum temperature higher than 32°F.

Unobserved heterogeneity caused by spatial and temporal correlations of data often can be found by visualizing the data and corroborated with statistical methods. The yearly average fatal crash frequencies for each county in Iowa is shown in Figure 1. As expected, fatal crash data revealed a cluster of high numbers of crashes in the central counties around the yellow-shaded area, where the largest city of Iowa, Des Moines, is located. Fatal crash data also revealed a cluster of low numbers of crashes in the northern and southwestern parts of Iowa (deep-shaded areas). Next, statistical analysis was performed to investigate the presence of spatial correlations.

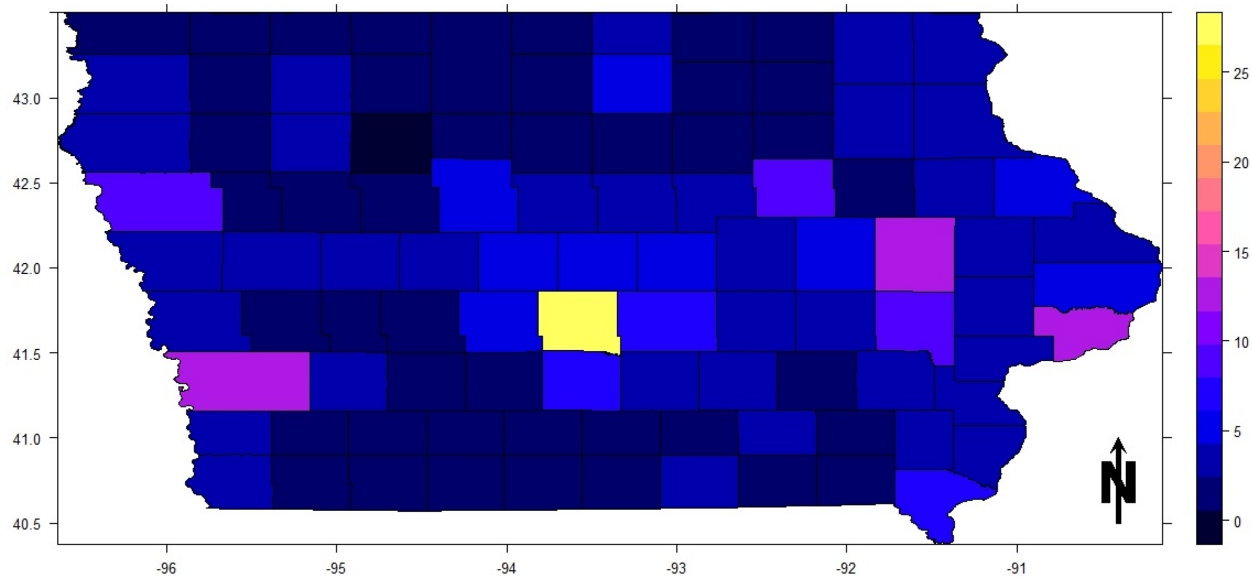


Figure 1 County-level yearly average fatal crash counts of Iowa (2006-2015)

Moran's *I* statistic is commonly used to test spatial correlations in traffic crash analysis (Guo et al., 2010; Quddus, 2008; Xie et al., 2014; Zeng and Huang, 2014). The global Moran's *I* is defined as (Anselin, 1988):

$$I = \frac{n\sum_i \sum_i \omega_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i \neq j} \omega_{ij} \sum_i (y_i - \bar{y})^2}, \tag{1}$$

where n is the total number of observations, y_i and y_j are the values of observation i and observation j, $\bar{y}$ is the average value of observations, and ω_{ij} is the spatial weight between observations i and j.

Negative Moran's *I* values indicate negative spatial autocorrelation, positive values indicate positive spatial autocorrelation, and zero indicates no spatial autocorrelation. The *z*-score of Moran's *I* shows if the spatial autocorrelation is significant.

The global Moran's *I* statistics of fatal crashes in each year from 2006 to 2015 were calculated using the "spdep" package (Bivand and Piras, 2015) in the R platform (R Core Team, 2016) with queen continuity spatial weights, whereby counties with a shared border or vertex were considered as neighbors. When areas were neighbors, the spatial weights were 1; otherwise, they were 0. The results are shown in Table 2.

Table 2 Global Moran's *I* statistics of fatal crashes in each year

Year	Moran's *I*	*P*-value
2006	1.986	0.024*
2007	2.091	0.018*
2008	1.520	0.064
2009	1.661	0.048*
2010	2.486	0.006*

2011	1.919	0.027*
2012	1.240	0.108
2013	2.387	0.008*
2014	1.241	0.107
2015	2.300	0.011*

Note: *significant at $P = 0.05$.

Significant spatial autocorrelations for fatal crashes existed in 7 out of 10 years at a 95% confidence level and at a 90% confidence level for the remaining 3 years. Thus, fatal crashes were highly likely to be spatially correlated at the county level in Iowa. These trends may be site specific. For example, Aguero-Valverde and Jovanis (2006) found the county-level yearly fatal crashes of Pennsylvania to not be significantly correlated. This suggests that the presence and type of spatial correlation is site and data sensitive. Therefore, no prior assumptions should be made about the presence or absence of spatial correlation, and it is recommended to statistically test the presence of spatial correlation prior to modeling.

Temporal correlation was not directly tested, as there were only 10 time points in this dataset. However, as shown in Figure 2, the yearly fatal crash counts of Iowa from 2006 to 2015 revealed a linearly decreasing trend, which needed to be considered when building the model.

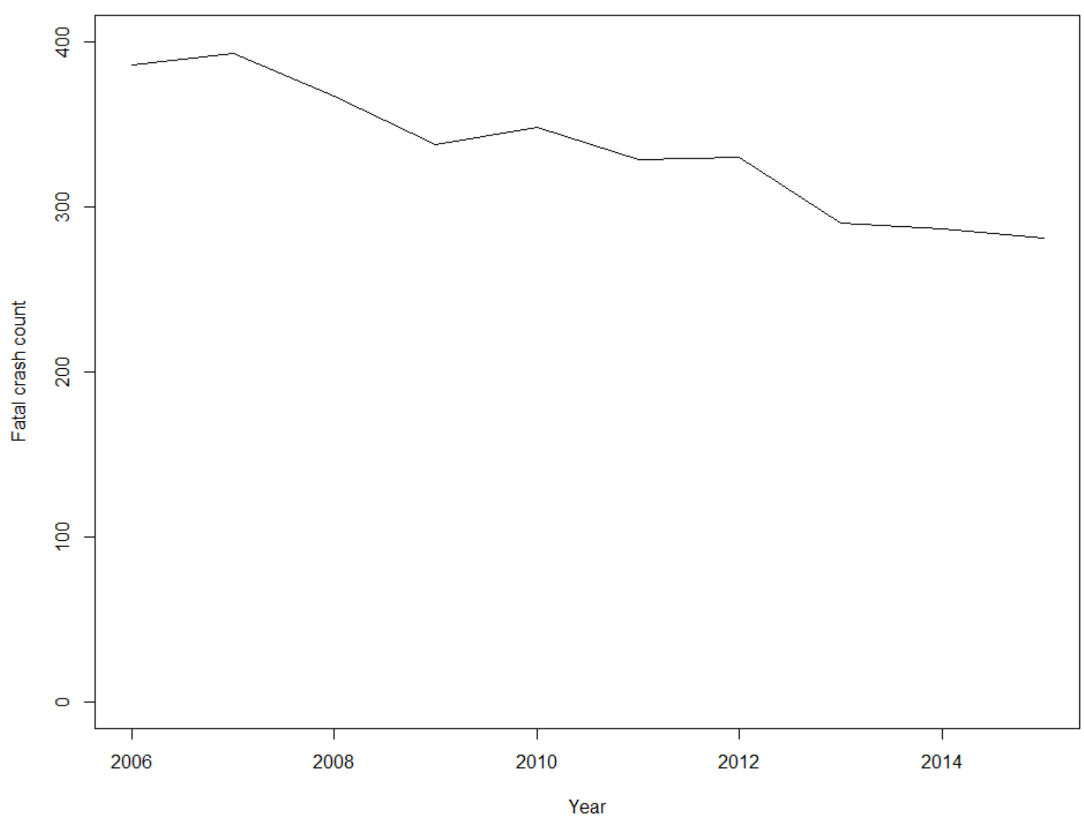


Figure 2 Iowa state-level yearly fatal crash counts (2006-2015)

3 Methodology

3.1 Statistical Framework

The statistical framework uses a Bayesian hierarchical architecture, including both the spatial and temporal random effect components. The statistical model is presented in equations 2 and 3:

$$y_{it} \sim Poisson(\lambda_{it}) \tag{2}$$

$$\log(\lambda_{it}) = \alpha + \beta * X_{it} + \upsilon_i + \nu_i + \varphi_t + \eta_{it}, \tag{3}$$

where i is the county number, 1,2,… 99; t is the year, 1 (2006), 2 (2007), …,10 (2015); y_{it} is the crash count of county i in year t; λ_{it} is the mean crash frequency of county i in year t; α is the intercept term; β is the regression coefficient vector; X_{it} is the covariate vector of county i in year t; υ_i is the structured spatial random effect of county i; ν_i is the unstructured spatial random effect of county i; φ_t is the temporal random effect in year t; and η_{it} is the spatio-temporal interaction effect.

The spatial and temporal components helped us to identify the underlying unobserved heterogeneity across county and year. For this study, we analyzed three kinds of spatio-temporal models that had the same spatial component but different temporal components.

3.1.1 Spatial Component

The spatial component, i.e. $\upsilon_i + \nu_i$, was assumed to follow the Besag-York-Mollie (BYM) model (Besag et al., 1991). The BYM model has been widely used in traffic accident analysis (Aguero-Valverde and Jovanis, 2006; Boulieri et al., 2017; Wang et al., 2013; Xie et al., 2014) and has been recommended for traffic crash analyses (Boulieri et al., 2017). For the BYM model, the structured spatial effect, υ_i, is modeled using an intrinsic conditional autoregressive (ICAR) structure, and the unstructured spatial effect, ν_i, follows a normal distribution.

$$\upsilon_i | \upsilon_{j \neq i} \sim N\left(\frac{\sum_{j \in N(i)} \upsilon_j}{\# N(i)}, \frac{{\tau_\upsilon}^{-1}}{\# N(i)}\right) \tag{4}$$

$$\nu_i \sim N(0, {\tau_\nu}^{-1}), \tag{5}$$

where $N(i)$ are the neighbors of county i, $\#N(i)$ are the number of neighbors of county i, and τ_υ and τ_ν are precisions.

The ICAR part accounts for possible spatial correlations between counties, and the unstructured part is responsible for county individual heterogeneity.

3.1.2 Temporal Component

Three temporal models, including the linear temporal model, the 1st order autoregressive (AR1) model, and the 1st order random walk (RW1) model, were considered.

The linear temporal model is defined in equations 6 and 7 (Bernardinelli et al., 1995):

$$\varphi_t = (\beta_2 + \delta_i) * t \tag{6}$$

$$\delta_i \overset{iid}{\sim} N(0, \tau_\delta^{-1}), \tag{7}$$

where β_2 is the global time trend; δ_i is the interaction between time and county i, $\delta_i < 0$ implies that the area-specific trend is smaller than the mean trend, whereas $\delta_i > 0$, implies that the area-specific trend is larger than the mean trend; and τ_δ is the precision.

δ_i could reflect the degree to which spatial effects and temporal effects have interactions (Blangiardo et al., 2013).

The AR1 model is defined in equations 8, 9, and 10:

$$\varphi_t \sim \begin{cases} N\left(0, \left(\tau_\varphi(1-\rho^2)\right)^{-1}\right) & for\ t = 1 \\ \rho\varphi_{t-1} + \varepsilon_t & for\ t = 2,3,..10 \end{cases} \tag{8}$$

$$|\rho| < 1 \tag{9}$$

$$\varepsilon_t \sim N(0, \tau_\varepsilon^{-1}), \tag{10}$$

where ρ is a correlation parameter, ε_t is the white noise, and τ_ε is a precision.

The RW1 model is defined in equations 11 and 12:

$$\varphi_{t+1} = \varphi_t + \gamma_t \tag{11}$$

$$\gamma_t \overset{iid}{\sim} N\left(0, \tau_\gamma^{-1}\right), \tag{12}$$

where γ_t is the white noise and τ_γ is a precision.

3.1.3 Spatio-Temporal Component

The spatio-temporal component, η_{it}, is assumed to follow a zero-mean normal distribution.

$$\eta_{it} \overset{iid}{\sim} N\left(0, \tau_\eta^{-1}\right). \tag{13}$$

where τ_η is a precision.

Due to the presence of η_{it}, this statistical model becomes the Poisson log-normal model.

In addition, the performance of the best spatio-temporal model, which is the linear temporal component model as proven later, is compared against several traditional models discussed below.

3.1.4 Other Comparison Models

3.1.4.1 Spatial Effects and Temporal Effects Assessment

Three models, one with no spatial or temporal effects, one with only spatial effects, and one with only temporal effects, were compared against the best spatio-temporal model to assess the importance of explicitly accounting for spatial and temporal effects.

3.1.4.2 Poisson Model vs. Zero-Inflated Poisson (ZIP) model

As shown in Table 1, fatal crashes had zero inflation. Thus, the ZIP model was also built for comparison. It should be noted that for zero-inflated crash data, the zero-state Markov switching model has been shown to be superior to the zero-inflated model (Malyshkina and Mannering, 2010). However, the zero-state Markov switching model is not discussed here, as the focus is on explaining zero inflation caused by spatial or temporal correlations and hence can be explicitly explained using a ZIP model.

All combinations of spatial, temporal, and base case models explored in this study are listed in Table 3.

Table 3 Summary of models developed for fatal crash frequency analysis

No	Model code	Spatial effect	Temporal effect	Spatio-temporal effect	Base model
1	$S_0T_0ST_0P$	—	—	—	Poisson
2	$S_{BYM}T_0ST_0P$	BYM	—	—	Poisson
3	$S_0T_LST_0P$	—	Linear	—	Poisson
4	$S_{BYM}T_LST_0P$	BYM	Linear	—	Poisson
5	$S_{BYM}T_LST_1P$	BYM	Linear	η_{it}	Poisson
6	$S_{BYM}T_{AR1}ST_1P$	BYM	AR1	η_{it}	Poisson
7	$S_{BYM}T_{RW1}ST_1P$	BYM	RW1	η_{it}	Poisson
8	$S_{BYM}T_LST_1ZIP$	BYM	Linear	η_{it}	ZIP

Note: 0, component not included; 1, component included; L, linear temporal; BYM, Besag-York-Mollie; AR1, 1st order autoregressive; RW1, 1st order random walk; ZIP, zero-inflated Poisson; "—" means non-existent.

3.2 Integrated Nested Laplace Approximation (INLA)

Bayesian models are usually solved with Markov chain Monte Carlo (MCMC) simulations. However, when the models are very complex without close-form posterior density available, as in this case, the MCMC method can be very time consuming if both spatial and temporal effects are included. Rue and Martino (2009) proposed the INLA method to numerically approximate the full Bayesian inference for latent Gaussian models. INLA can produce much faster results than can the MCMC approach for Bayesian models without compromising accuracy (Martins et al., 2013), as it can accurately derive the posterior densities by numerical approximation and significantly decrease the MCMC simulation workload.

Assume y is the response vector, θ is the target parameter vector, and ψ is the hyper-parameter vector. The posterior probability densities of parameter elements and hyper-parameter elements in Bayesian models are (Blangiardo et al., 2013):

$$p(\theta_i|y) = \int p(\psi|y)p(\theta_i|\psi, y)d\psi \quad (14)$$

$$p(\psi_k|y) = \int p(\psi|y)d\psi_{-k}, \quad (15)$$

where i is the ith observation; θ_i is the ith parameter; ψ_k is the kth hyper-parameter; and ψ_{-k} is the complement hyper-parameter set to ψ_k.

The INLAs for the posterior densities of interest can be written as (Blangiardo et al., 2013; Rue et al., 2009):

$$p(\psi|y) = \frac{p(\theta,\psi|y)}{p(\theta|\psi,y)} \propto \frac{p(\psi)p(\theta|\psi)p(y|\theta)}{p(\theta|\psi,y)} \approx \frac{p(\psi)p(\theta|\psi)p(y|\theta)}{\tilde{p}(\theta|\psi,y)}|_{\theta=\theta^*(\psi)} =: \tilde{p}(\psi|y) \tag{16}$$

$$p(\theta_i|\psi,y) = \frac{p((\theta_i,\theta_{-i})|\psi,y)}{p(\theta_{-i}|\theta_i,\psi,y)} \approx \frac{p(\theta,\psi|y)}{\tilde{p}(\theta_{-i}|\theta_i,\psi,y)}|_{\theta_{-i}=\theta^*{}_{-i}(\theta_i,\psi)} =: \tilde{p}(\theta_i|\psi,y), \tag{17}$$

where $\tilde{p}(\psi|y)$ is the Gaussian approximation of $p(\theta|\psi,y)$ and $\theta^*(\psi)$ is its mode and $\tilde{p}(\theta_{-i}|\theta_i,\psi,y)$ is the simplified Laplace approximation based on the Taylor's series expansion of the Laplace approximation of $\tilde{p}(\theta_i|\psi,y)$.

As compared to the Gaussian approximation, the simplified Laplace approximation in equation 17 provides a good balance between speed and accuracy.

INLA first obtains the marginal joint posterior of $\tilde{p}(\psi|y)$ to locate the mode by grid search. Then, for each ψ^* with the corresponding weight $w_{\psi}{}^*$, the conditional posteriors $\tilde{p}(\theta_i|\psi^*,y)$ are also obtained by grid search. Finally, the marginal posteriors $\tilde{p}(\theta_i|y)$ are obtained by numerical integration:

$$\tilde{p}(\theta_i|y) \approx \sum_{\psi^* \in G} \tilde{p}(\theta_i|\psi^*,y)\tilde{p}(\psi^*|y)w_{\psi}{}^*. \tag{18}$$

More details about INLA can be found elsewhere (Blangiardo et al., 2013; Hu et al., 2013; Martins et al., 2013; Rue et al., 2009).

All eight models listed in Table 3 were implemented in the R environment (R Core Team, 2016) using the 'INLA' package (Lindgren and Rue, 2015; Martins et al., 2013; Rue et al., 2009). The regression coefficients β were assigned independent normal distributions $N(0, 1000)$. Six hyper-parameters are defined in this study, i.e. the precision parameters τ_υ, τ_ν, τ_δ, τ_ε, τ_γ, and τ_η. The logarithm of these values were assigned to follow the log-Gamma distribution $logGamma(1{,}0.0005)$ (Blangiardo et al., 2013).

3.3 Model Comparison and Checking

The deviance information criterion (DIC) was used as a measure of assessing different Bayesian models (Spiegelhalter et al., 2002). DIC is defined as

$$DIC = D(\bar{\theta}) + 2p_D = \overline{D} + p_D, \tag{19}$$

where $D(\bar{\theta})$ is the deviance using the posterior mean values of the estimated parameters ($\bar{\theta}$), $\overline{D}$ is the posterior mean of deviances, and p_D is the effective number of parameters.

Similar to Akaike's information criterion (AIC), DIC considers both the Bayesian measure of fit or adequacy and the complexity of the model (Spiegelhalter et al., 2002). Models with smaller DIC values are expected to perform better. Roughly, differences of more than 10 might definitely rule out the model with the higher DIC, differences between 5 and 10 are substantial, and differences less than 5 might mean that the models are not significantly different (MRC Biostatistics Unit, 2004).

However, DIC may under-penalize complex models with many random effects (Plummer, 2008), such as CAR models. Thus, the conditional predictive ordinate (CPO) (Pettit, 1990) and the cross-validated probability integral transform (PIT) (Dawid, 1984) were also calculated for model assessment. Both of them are leave-one-out cross validation scores.

$$CPO_i = \pi(y_i|\boldsymbol{y}_{-i}) \tag{20}$$

$$PIT_i = p(Y_i \leq y_i|\boldsymbol{y}_{-i}), \tag{21}$$

where ${y_i}^{obs}$ is the ith observation and $\boldsymbol{y}_{-i}$ represents all the observations except the ith one.

The negative mean logarithmic CPO was calculated as a measure of the predictive quality of the model (Gneiting and Raftery, 2007; Roos and Held, 2011).

$$\overline{CPO} = -\frac{1}{n}\sum_i^n \log(CPO_i) \tag{22}$$

Stone (1977) proved that the $\overline{CPO}$ was asymptotically equivalent to AIC. Thus, $\overline{CPO}$ can be used for model choice, and a lower value of $\overline{CPO}$ indicates a better model.

A large or small PIT value indicates possible outliers, and the PIT values of a well-calibrated model should be uniformly distributed. Thus PIT histograms can be used to assess the calibration of a model (Czado et al., 2009). For count data, an adjusted PIT should be used instead to make the predictive distribution continuous (Czado et al., 2009).

$$Adjusted\ PIT_i = PIT_i - \frac{1}{2}CPO_i \tag{23}$$

In addition, root mean square error (RMSE) and mean absolute error (MAE) were also calculated to evaluate the adequacy of model fit.

$$RMSE = \sqrt{\frac{1}{n_0}\sum_{j=1}^{n_0}\left(O_j - P_j\right)^2} \tag{24}$$

$$MAE = \frac{1}{n_0}\sum_{j=1}^{n_0}\left|O_j - P_j\right|, \tag{25}$$

where O_j is the jth observation value, P_j is the predicted ith value from the model, and n_0 is the number of observations.

Similar to DIC, smaller MAE and RMSE values are desired.

3.4 Spatial Fraction Analysis

For the spatio-temporal analysis, one point of interest was to identify the contribution of the structured spatial effects ${\sigma_\upsilon}^2$ over the total marginal spatial variability ${\sigma_\upsilon}^2 + {\sigma_\nu}^2$ (Boulieri et al., 2017). The spatial fraction of interest is given by

$$frac_\upsilon = \frac{{\sigma_\upsilon}^2}{{\sigma_\upsilon}^2 + {\sigma_\nu}^2} = \frac{1/\tau_\upsilon}{1/\tau_\upsilon + 1/\tau_\nu}, \tag{26}$$

where ${\sigma_\upsilon}^2$ is the variance of the structured spatial effects, ${\sigma_\nu}^2$ is the variance of the unstructured spatial effects, and τ_υ and τ_ν are the corresponding precisions.

When the spatial fraction is close to 1, the structured spatial effects explain most of the variability of the model. Otherwise, the unstructured spatial random effects play the main role.

4 Results and Discussions

All eight models listed in Table 3 were implemented in INLA. On an Intel(R) Xeon(R) CPU at 3.70 GHz with 16 GB random access memory, it took a total of 73.074 sec to run these eight models. As a comparison, it took INLA 13.609 sec to estimate the $S_{BYM}T_{L}ST_{1}P$ model, whereas it took OpenBUGS (Sturtz et al., 2005) 1,053 sec to estimate the same model with the MCMC simulation settings of three simulation chains, 5,000 burn-in samples, and 5,000 adopted samples with a thin interval set at 2. The computation time was greatly reduced using INLA, and the computation time is expected to be saved more with the increase of data and parameters.

The DIC, $\overline{CPO}$, RMSE, and MAE values of the eight models listed in Table 3 are shown in Table 4. These four measures help in identifying the best spatio-temporal model. The following observations can be made from data shown in Table 4.

Table 4 DIC, $\overline{CPO}$, and RMSE, MAE values for all the models

No	Model	DIC	$\overline{CPO}$	RMSE	MAE
1	$S_0T_0ST_0P$	4282.01	2.172	2.652	1.810
2	$S_{BYM}T_0ST_0P$	3791.62	1.920	1.851	1.350
3	$S_0T_LST_0P$	3860.00	1.953	1.987	1.421
4	$S_{BYM}T_LST_0P$	3749.39	1.896	1.757	1.315
5	$S_{BYM}T_LST_1P$	3746.13	1.894	1.757	1.314
6	$S_{BYM}T_{AR1}ST_1P$	3750.60	1.899	1.762	1.316
7	$S_{BYM}T_{RW1}ST_1P$	3752.33	1.896	1.765	1.319
8	$S_{BYM}T_LST_1ZIP$	3749.35	1.895	1.756	1.314

Note: 0, component not included; 1, component included; L, linear temporal component; BYM, Besag-York-Mollie; AR1, 1st order autoregressive; RW1, 1st order random walk; ZIP, zero-inflated Poisson; "—" means non-existent.

4.1 Choice of the Temporal Component

The DIC values do not show significant differences among the $S_{BYM}T_{L}ST_{1}P$, $S_{BYM}T_{AR1}ST_{1}P$, and $S_{BYM}T_{RW1}ST_{1}P$ models, but the $S_{BYM}T_{L}ST_{1}P$ model with the linear temporal component had the lowest $\overline{CPO}$, RMSE, and MAE values. In addition, the adjusted PIT histogram of the $S_{BYM}T_{L}ST_{1}P$ model is shown in Figure 3, where the adjusted PIT values show a very good uniform distribution. This implies that the $S_{BYM}T_{L}ST_{1}P$ model was well calibrated for the data. Thus, the $S_{BYM}T_{\mathrm{L}}ST_{1}P$ model was considered as the best fit in this case; that is, fatal crash frequencies had some linear changing trend. Although these models did not show large differences, the results still implied the necessity of temporal component selection, especially considering that different temporal models would lead to different interpretations of the data. For example, the linear temporal component implies that the number of fatal crashes would change linearly in the future, but the same conclusion may not be drawn from the RW1 temporal component.

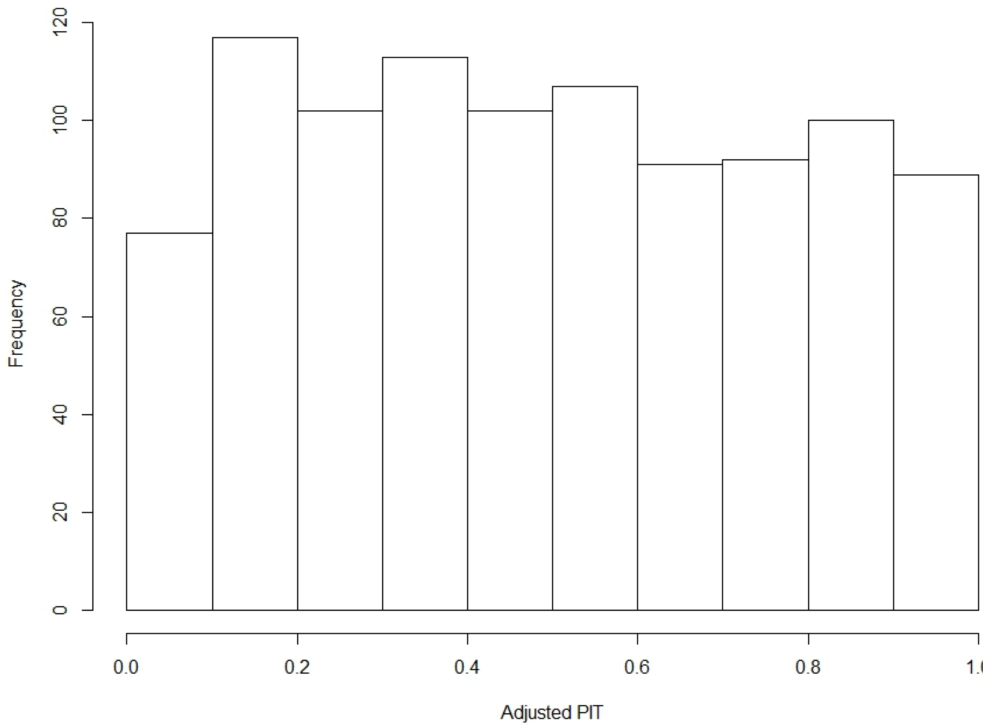


Figure 3 Histogram of the adjusted PIT values of the $S_{BYM}T_LST_1P$ model

4.2 Necessity of Including Spatial, Temporal, and Spatio-Temporal Effects

The $S_{BYM}T_LST_0P$ model performed much better than did the $S_0T_0ST_0P$, $S_{BYM}T_0ST_0$, and $S_0T_LST_0P$ models in terms of all four measures. This means that, in this case, both spatial and temporal effects played important roles in unobserved heterogeneity and thus needed to be considered. Meanwhile, because the $S_{BYM}T_0ST_0P$ model had much lower DIC, $\overline{CPO}$, RMSE, and MAE values than did the $S_0T_LST_0P$ model, spatial effects had a greater influence than did temporal effects in this case. This finding indicates that fatal crashes have very strong correlations across counties in Iowa. Only 10 years of data were used for this study, and it may not be a long enough time span for crashes to show a big change over time. If more years of data were available or monthly data had been analyzed, the temporal effects may have played a more important role. The $S_{BYM}T_LST_1P$ model was slightly better than $S_{BYM}T_LST_0P$ model, which meant the spatio-temporal interaction effects were very weak.

4.3 Zero-Inflation of Crashes

The $S_{BYM}T_LST_1P$ model had nearly the same performance as the $S_{BYM}T_LST_1ZIP$ model did in terms of all four measures. In addition, the zero-inflation probability value, which showed the probability of zero crashes being from the zero state, was only 0.0046 for the $S_{BYM}T_LST_1ZIP$ model. This means that there was no longer a need to consider zero inflation after including spatial and temporal effects, as the zero inflation of fatal crashes could be well explained by spatial and temporal effects. This finding provides a new point of view for the explanation of where zero inflation comes from in crash data.

Because the $S_{BYM}T_LST_1P$ model had the best performance of all eight models, it was used in the following analysis. The estimated parameters, their standard errors, and 95% credible intervals are shown in Table 5. As expected, VMT had significant positive effects. However, all the other variables were statistically insignificant. It is thought that population, employment rate, and income indicators in Iowa had been relatively consistent from 2006 to 2015 because Iowa was a typical farming state and there were no significant changes in these variables. Thus, these indicators did not show significant influences. In addition, although adverse weather may increase the number of crashes in the short term, the results here show that weather may not have a big influence on fatal crashes in the long term in Iowa.

Table 5 Estimated parameters of the $S_{BYM}T_LST_1P$ model with all covariates

Parameter	Mean	Std. Err.	0.025 quantile	0.975 quantile
(Intercept)	0.427	0.431	–0.420	1.272
VMT	0.887	0.082	0.727	1.049
Population	–0.003	0.003	–0.010	0.003
Income	–0.014	0.036	–0.085	0.058
Unemployment rate	0.013	0.020	–0.027	0.053
Rainfall	–0.002	0.003	–0.007	0.003
Snowfall	0.000	0.002	–0.003	0.003
TH32	0.002	0.002	–0.001	0.006
Year	–0.041	0.006	–0.053	–0.029

Note: VMT, vehicle miles traveled; TH32, number of days with minimum temperature higher than 32°F.

Because only the VMT parameter was significant, the $S_{BYM}T_LST_1P$ model was rebuilt using only VMT. The results are shown in Table 6.

Table 6 Estimated parameters of the $S_{BYM}T_LST_1P$ model with only VMT

	Intercept	VMT (β_1)	Year (β_2)	τ_υ	τ_ν	τ_δ	$frac_\upsilon$
Mean	0.923	0.887	–0.042	9.919	9.812	16424.166	0.497
Std. Err.	0.057	0.086	0.006	7.298	3.446	12860.000	—
0.025 quantile	0.810	0.714	–0.054	2.692	4.807	1926.189	—
0.975 quantile	1.032	1.046	–0.030	31.290	16.040	55533.450	—

Note: VMT, vehicle miles traveled; β_1, β_2, regression coefficients; $\tau_\upsilon, \tau_\nu, \tau_\delta$, precisions; $frac_\upsilon$ = spatial fraction.

4.4 Spatial Fraction Results

For the $S_{BYM}T_LST_1P$ model, the fraction of structured spatial effects was 0.497 (Table 6), which implied that the unstructured and structured spatial effects played nearly the same role in this case. That is, the unobserved heterogeneity in space existed both between counties and for individual counties. The exponential posterior means of the structured spatial effects of each county were shown in Figure 4; the counties with $\exp(\upsilon_i)$ lower than 1 tended to have fewer crashes and the counties with $\exp(\upsilon_i)$ greater than 1 tended to have more crashes. As shown in Figure 4, the counties located in northern and southwestern Iowa tended to have fewer fatal crashes. This finding is generally consistent with the empirically observed fatal crash distribution shown in Figure 1.

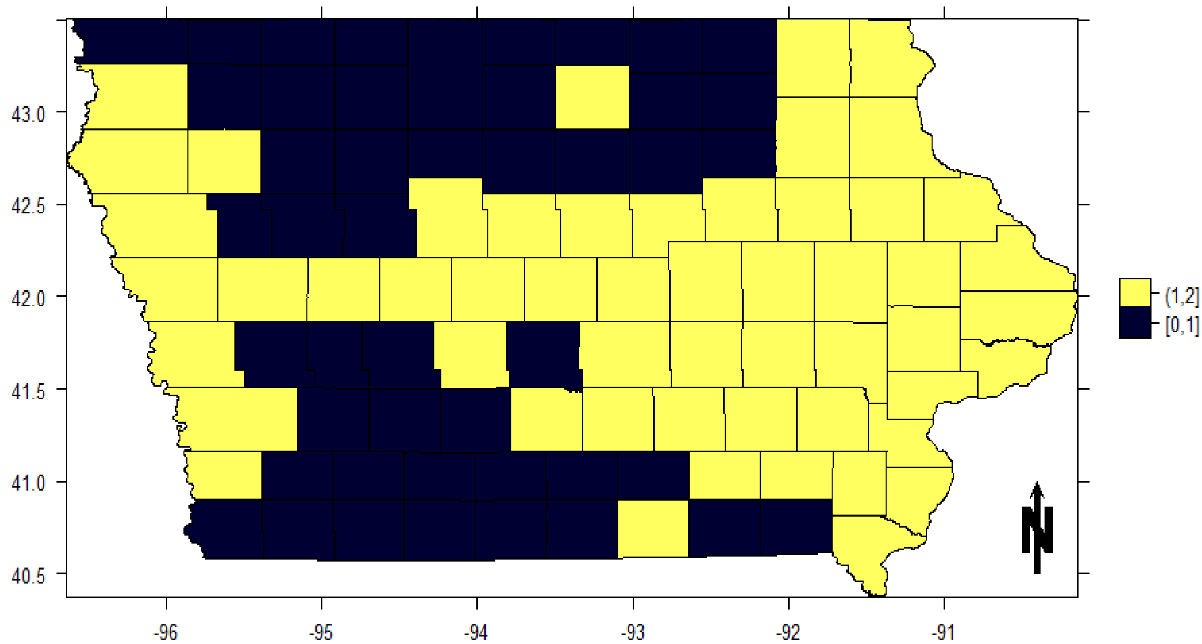


Figure 4 Exponential posterior means of the structured spatial effect ($\boldsymbol{exp}(\upsilon_i)$)

Moran's I statistics of the residuals of the $S_{BYM}T_LST_1P$ model were calculated to see if they still had spatial correlations. As shown in Table 7, the p-values of residuals were significantly larger than 0.05 for any year except 2010, the p-value of which was very close to 0.05. Thus, the spatial component covered nearly all of unobserved heterogeneity in space. The results also verified the effectiveness of the spatial component.

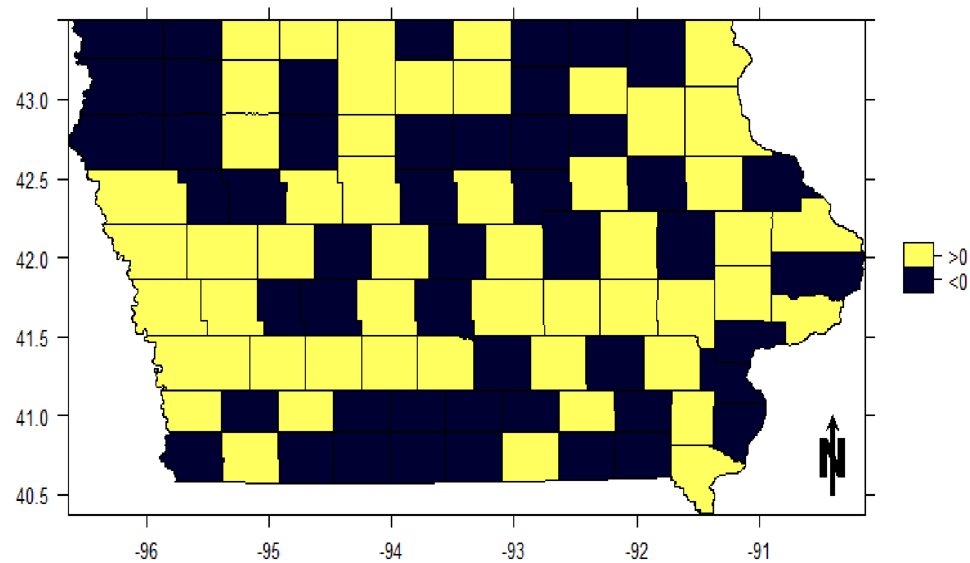
Table 7 Moran's *I* test results for the residuals of the $\boldsymbol{S_{BYM}T_LST_1P}$ model

Year	Moran's *I* statistic	*p*-value
2006	–1.036	0.850
2007	–0.156	0.562
2008	–0.792	0.786
2009	–0.535	0.704
2010	1.653	0.049

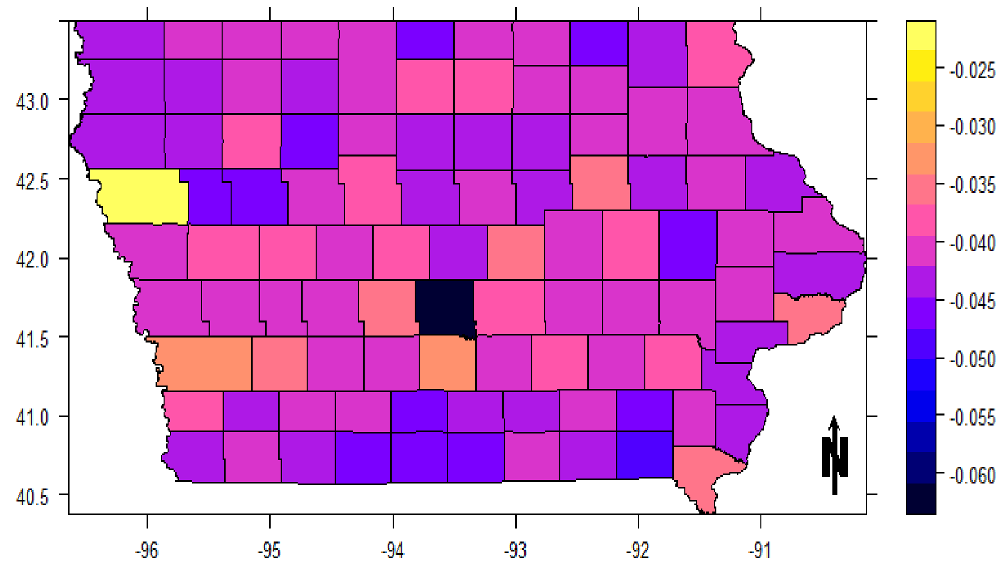
2011	0.292	0.385
2012	0.636	0.262
2013	0.460	0.323
2014	–0.876	0.809
2015	–1.387	0.917

4.5 Temporal Effects

The β_2 value of –0.042 with a 95% credible interval of [–0.054, –0.030] means that, on average, fatal crashes in Iowa significantly decreased from 2006 to 2015. The signs of δ_i values, a positive value meaning that the number of fatal crashes of county i decreased slower than the state average value and a negative value meaning that the number of fatal crashes of county i decreased faster than the state average value, are shown in Figure 5(a). The changing rates of fatal crashes for each county, i.e. $\beta_2 + \delta_i$, are shown in Figure 5(b). All $\beta_2 + \delta_i$ values were negative, which meant that the number of fatal crashes for all the counties showed decreasing trends from 2006 to 2015. The δ_i values for 50 out of the total of 99 counties were positive, whereas the δ_i values for the remaining 49 counties were negative; that is, the number of fatal crashes in 50 counties decreased slower than the mean trend of the whole state, whereas fatal crash numbers in the remaining 49 counties decreased faster than the mean trend. Thus, the first 50 counties should be the focus of future traffic safety improvement programs.



(a) δ_i



(b) $\beta_2+\delta_i$

Figure 5 Iowa county-level fatal crash yearly change trends from 2006 to 2015

5 Conclusions and Future Research

Unobserved heterogeneity due to the correlations of crashes in space and time has been proven to be a big issue in many studies. However, only a limited number of studies have considered both of them in modeling crash frequency. This study explored spatial and temporal effects in crash frequency models to account for unobserved heterogeneity and accurately identified the long-term regional trends in the change of traffic crash frequencies. Focusing on the number of yearly fatal crashes at the county level in Iowa from 2006 to 2015, multiple spatio-temporal models with the same spatial component but different temporal components were developed using the Bayesian framework. INLA, a fast Bayesian model estimation methodology, was used to estimate parameters. The model with a linear temporal component was found to be the most appropriate. Numbers of fatal crashes in all Iowa counties were found to show linearly decreasing trends but with different rates of decrease by counties. No explanatory factors, except VMT, were found to have a significant influence on fatal crash frequencies. Spatial and temporal effects were found to be responsible for both over dispersion and zero inflation of crash data, whereas spatial effects played a more important role than did temporal effects in this case.

In future research, the impact of a smaller time scale, such as season or month, should be explored, as this may offer more details about crash frequency changing trends and show the influences of periodic factors such as weather. Meanwhile, although zero inflation is not a problem any more with the use of the spatio-temporal model for this dataset, this may not be true for other datasets. When the spatio-temporal model does not explain excess zeros completely, the zero-state Markov switching model may be combined with spatial effects to develop new spatio-temporal models. The zero-state Markov switching model could account for both zero inflation and temporal correlations, and it has been proven to be superior to traditional zero-inflated models (Malyshkina and Mannering, 2010). Finally, as Boulieri et al. (2017) has suggested, the multivariate space–time model considering factorial space and time interactions can be evaluated to better exploit spatial, temporal, and between-variable correlations, but this may need high performance computing along with complex modeling structure.

6 References

Abdel-Aty, M., Wang, X., 2006. Crash estimation at signalized intersections along corridors: analyzing spatial effect and identifying significant factors. Transportation Research Record 1953, 98–111.

Aguero-Valverde, J., 2011. Direct spatial correlation in crash frequency models. 3rd International Conference on Road Safety and Simulation, Indianapolis, IN, USA.

Aguero-Valverde, J., Jovanis, P.P., 2008. Analysis of road crash frequency with spatial models. Transportation Research Record 2061, 55–63.

Aguero-Valverde, J., Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. Accident Analysis and Prevention 38 (3), 618–625.

Ahmed, M.M., Abdel-Aty, M., 2015. Evaluation and spatial analysis of automated red-light running enforcement cameras. Transportation Research Part C 50, 130–140.

Andrey, J., 2010. Long-term trends in weather-related crash risks. Journal of Transport Geography 18 (2), 247–258.

Anselin, L., 1988. Spatial Econometrics: Methods And Models. Springer, Netherlands.

Bernardinelli, L., Clayton, D., Pascutto, C., Montomoli, C., Ghislandi, M., Songini, M., 1995. Bayesian analysis of space-time variation in disease risk. Statistics in Medicine 14 (21–22), 2433–2443.

Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society Series B 36 (2), 192–263.

Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration, with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics 43 (1), 1–20.

Bivand, R., Piras, G., 2015. Comparing implementations of estimation methods for spatial econometrics. Journal of Statistical Software 63 (18), 1–36.

Blangiardo, M., Cameletti, M., Baio, G., Rue, H., 2013. Spatial and spatio-temporal models with R-INLA. Spatial and Spatio-temporal Epidemiology 7, 39–55.

Boulieri, A., Liverani, S., Hoogh, K. de, Blangiardo, M., 2017. A space–time multivariate Bayesian model to analyse road traffic accidents by severity. Journal of the Royal Statistical Society Series A 180 (1), 119–139.

Brijs, T., Karlis, D., Wets, G., 2008. Studying the effect of weather conditions on daily crash counts using a discrete time-series model. Accident Analysis and Prevention 40 (3), 1180–1190.

Czado, C., Gneiting, T., Held, L., 2009. Predictive model assessment for count data. Biometrics 65 (4), 1254–1261.

Dawid, A.P., 1984. Present position and potential developments: some personal views: statistical theory: the prequential approach. Journal of the Royal Statistical Society Series A 147 (2), 278–292.

Eckley, D.C., Curtin, K.M., 2013. Evaluating the spatiotemporal clustering of traffic incidents. Computers, Environment and Urban Systems 37, 70–81.

El-Basyouny, K., Barua, S., Islam, M.T., 2014. Investigation of time and weather effects on crash types using full Bayesian multivariate Poisson lognormal models. Accident Analysis and Prevention 73, 91–99.

Erdogan, S., 2009. Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey. Journal of Safety Research 40 (5), 341–351.

Gneiting, T., Raftery, A.E., 2007. Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association 102 (477), 359–378.

Golob, T.F., Recker, W.W., 2003. Relationships among urban freeway accidents, traffic flow, weather, and lighting conditions. Journal of Transportation Engineering 129 (4), 342–353.

Guo, F., Wang, X., Abdel-Aty, M.A., 2010. Modeling signalized intersection safety with corridor-level spatial correlations. Accident Analysis and Prevention 42 (1), 84–92.

Hu, S., Ivan, J.N., Ravishanker, N., Mooradian, J., 2013. Temporal modeling of highway crash counts for senior and non-senior drivers. Accident Analysis and Prevention 50, 1003–1013.

Iowa Community Indicators Program, 2016. Iowa Community Indicators Program (ICIP). http://www.icip.iastate.edu/ (accessed 2.20.17).

Iowa Department of Transportation, 2016. Vehicle-miles traveled (VMT). http://www.iowadot.gov/maps/msp/vmt/vmt.html (accessed 10.20.16).

Iowa Environmental Mesonet, 2017. Iowa Environmental Mesonet (IEM). https://mesonet.agron.iastate.edu/ (accessed 2.20.17).

Jiang, X., Abdel-Aty, M., Alamili, S., 2014. Application of Poisson random effect models for highway network screening. Accident Analysis and Prevention 63, 74–82.

Kilamanua, W., Xia, J., Caulfieldb, C., 2011. Analysis of spatial and temporal distribution of single and multiple vehicle crash in Western Australia: a comparison study. 19th International Congress on Modelling and Simulation, Perth, Australia.

Knapp, K.K., Smithson, L.D., Khattak, A.J., 2000. The mobility and safety impacts of winter storm events in a freeway environment. Mid-Continent Transoprtation Symposium, Ames, IA, USA.

Levine, N.E.D., Kim, K.E., Nitz, L.H., 1995. Spatial analysis of Honolulu motor vehicle crashes: I. spatial patterns. Accident Analysis and Prevention 27 (5), 663–674.

Lindgren, F., Rue, H. avard, 2015. Bayesian spatial modelling with R-INLA. Journal of Statistical Software 63 (19), 1–26.

Liu, C., Gyawali, S., Sharma, A., Smaglik, E., 2015. A methodological approach for spatial and temporal analysis of red light running citations and crashes: a case-study in Lincoln, Nebraska. Transportation Research Board 94th Annual Meeting, Washington, D.C., USA.

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transportation Research Part A 44 (5), 291–305.

Lord, D., Washington, S.P., Ivan, J.N., 2005. Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Accident Analysis and Prevention 37 (1), 35–46.

Malyshkina, N. V., Mannering, F.L., 2010. Zero-state Markov switching count-data models: an empirical assessment. Accident Analysis and Prevention 42 (1), 122–130.

Malyshkina, N. V, Mannering, F.L., Tarko, A.P., 2009. Markov switching negative binomial models: an application to vehicle accident frequencies. Accident Analysis and Prevention 41 (2), 217–226.

Mannering, F.L., Bhat, C.R., 2014. Analytic methods in accident research: methodological frontier and future directions. Analytic Methods in Accident Research 1, 1–22.

Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Analytic Methods in Accident Research 11, 1–16.

Martins, T.G., Simpson, D., Lindgren, F., Rue, H., 2013. Bayesian computing with INLA: new features. Computational Statistics and Data Analysis 67, 68–83.

Matkan, A., Mohaymany, A., 2013. Detecting the spatial–temporal autocorrelation among crash frequencies in urban areas. Canadian Journal of Civil Engineering 40 (3), 195–203.

Maze, T.H., Agarwal, M., Burchett, G., 2005. Whether weather matters to traffic demand, traffic safety, and traffic flow. Transportation Research Record 1948, 170–176.

Miaou, S.-P., Song, J.J., Mallick, B.K., 2003. Roadway traffic crash mapping a space-time modeling approach. Journal of Transportation and Statistics 6 (1), 33–57.

Mitra, S., Washington, S., Schalkwyk, V., 2007. Important omitted spatial variables in safety models: understanding contributing crash causes at intersections. Transportation Research Board 86th Annual Meeting, Washington, D.C., USA.

MRC Biostatistics Unit, 2004. DIC: Deviance Information Criteria. https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-dic/ (accessed 8.23.17).

Mullahy, J., 1997. Heterogeneity, excess zeros, and the structure of count data models. Journal of Applied Econometrics 12 (3), 337–350.

Noland, R.B., Klein, N.J., Tulach, N.K., 2013. Do lower income areas have more pedestrian casualties? Accident Analysis and Prevention 59, 337–345.

Pettit, L.I., 1990. The conditional predictive ordinate for the normal distribution. Journal of the Royal Statistical Society Series B 52 (1), 175–184.

Plummer, M., 2008. Penalized loss functions for Bayesian model comparison. Biostatistics 9 (3), 523–539.

Ponicki, W.R., Gruenewald, P.J., Remer, L.G., 2013. Spatial panel analyses of alcohol outlets and motor vehicle crashes in California: 1999-2008. Accident Analysis and Prevention 55, 135–143.

Pulugurtha, S.S., Sambhara, V.R., 2011. Pedestrian crash estimation models for signalized intersections. Accident Analysis and Prevention 43 (1), 439–446.

Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: an analysis of London crash data. Accident Analysis and Prevention 40 (4), 1486–1497.

Quddus, M.A., 2008. Time series count data models: an empirical application to traffic accidents. Accident Analysis and Prevention 40 (5), 1732–1741.

R Core Team, 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Roos, M., Held, L., 2011. Sensitivity analysis in Bayesian generalized linear mixed models for binary data. Bayesian Analysis 6 (2), 259–278.

Rue, H., Martino, S., Chopin, N., 2009. Appoximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society Series B 71 (2), 319–392.

Savolainen, P.T., Mannering, F.L., Lord, D., Quddus, M.A., 2011. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accident Analysis and Prevention 43 (5), 1666–1676.

Song, J.J., Ghosh, M., Miaou, S., Mallick, B., 2006. Bayesian multivariate spatial models for roadway traffic crash mapping. Journal of Multivariate Analysis 97 (1), 246–273.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society Series B 64 (4), 583–639.

Stone, M., 1977. An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. Journal of the Royal Statistical Society Series B 39 (1), 44–47.

Sturtz, S., Ligges, U., Gelman, A., 2005. R2OpenBUGS: a package for running OpenBUGS from R. Journal of Statistical Software 12 (3), 1–16.

Sukhai, A., Jones, A.P., Love, B.S., Haynes, R., 2011. Temporal variations in road traffic fatalities in South Africa. Accident Analysis and Prevention 43 (1), 421–428.

Truong, L.T., Kieu, L.-M., Vu, T.A., 2016. Spatiotemporal and random parameter panel data models of traffic crash fatalities in Vietnam. Accident Analysis and Prevention 94, 153–161.

U.S. Bureau of Economic Analysis, 2016. Personal income sumamry: personal income, population, per capita personal income. https://www.bea.gov/iTable/iTable.cfm?reqid=70&step=30&isuri=1&7022=20&7023=7&7024=non-industry&7033=-1&7025=4&7026=xx,19000&7027=2015,2014,2013,2012,2011,2010&7001=720&7028=3&7030=0&7031=19000&7040=-1&7083=levels&7029=20&7090=70#reqid=70&step=30&isuri=1& (accessed 2.20.17).

Wang, C., Quddus, M.A., Ison, S.G., 2011. Predicting accident frequency at their severity levels and its application in site ranking using a two-stage mixed multivariate model. Accident Analysis and Prevention 43 (6), 1979–1990.

Wang, C., Quddus, M.A., Ison, S.G., 2009. Impact of traffic congestion on road accidents: a spatial analysis of the M25 motorway in England. Accident Analysis and Prevention 41 (4), 798–808.

Wang, C., Quddus, M., Ison, S., 2013. A spatio-temporal analysis of the impact of congestion on traffic safety on major roads in the UK. Transportmetrica A 9 (2), 124–148.

Wang, X., Abdel-Aty, M., 2006. Temporal and spatial analyses of rear-end crashes at signalized intersections. Accident Analysis and Prevention 38 (6), 1137–1150.

Wang, Y., Kockelman, K.M., 2013. A Poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. Accident Analysis and Prevention 60, 71–84.

Xie, K., Wang, X., Ozbay, K., Yang, H., 2014. Crash frequency modeling for signalized intersections in a high-density urban road network. Analytic Methods in Accident Research 2, 39–51.

Yannis, G., Antoniou, C., Papadimitriou, E., 2011. Autoregressive nonlinear time-series modeling of traffic fatalities in Europe. European Transport Research Review 3 (3), 113–127.

Zeng, Q., Huang, H., 2014. Bayesian spatial joint modeling of traffic crashes on an urban road network. Accident Analysis and Prevention 67, 105–112.